

Análisis Inteligente de Datos: Segundo Parcial

Claudio Sebastián Castillo

10 de mayo de 2022

ANOVA

Datos

Observaciones por grupo:

Se cumplen los supuestos para su implementación?

Anova

fit del modelo

coeficientes

p-value

F-value

Plot ANOVA

Conclusión

Testear homoscedasticidad

Test de Bartlett

sensibilidad al supuesto de normalidad

Testear normalidad

Testear normalidad analizando residuos

Anova y después: post-hoc

Tukey's Honest Significant Differences (HSD)

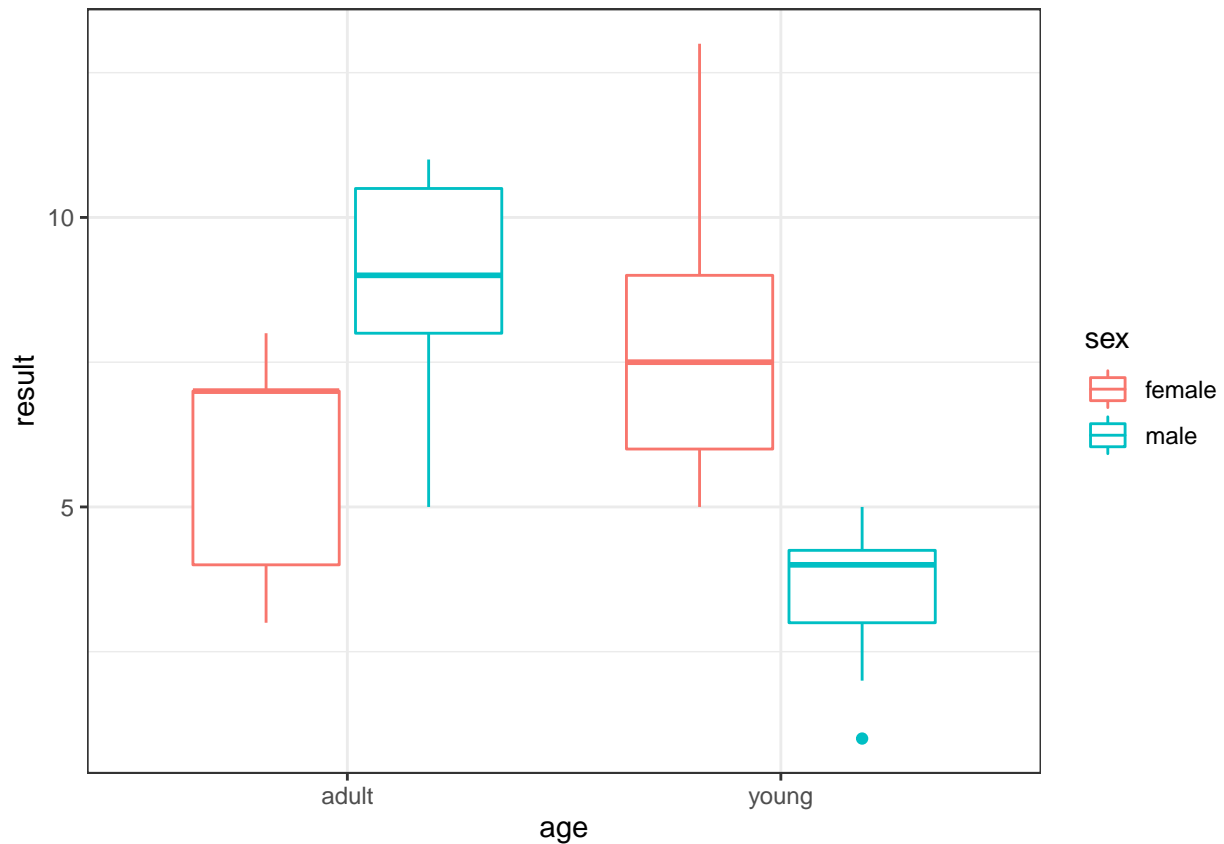
Cuando ANOVA no funciona: test de Kruskal-Wallis

ANOVA__multivariante

Datos

```
##      sex  age result
## 1 female adult     7
## 2  male adult    11
## 3  male adult     5
## 4 female adult     8
```

Gafico



Test Anovam

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## sex         1  15.13    15.13   3.691  0.06573 .
## age         1  41.95    41.95  10.233  0.00361 **
## sex:age      1  94.22    94.22  22.986 0.0000578 ***
## Residuals   26 106.57     4.10
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## [1] "La varianza presenta diferencias significativas entre los grupos"
```

Tamaño del efecto

.01: Small effect size .06: Medium effect size .14 or higher: Large effect size

```
##           eta.sq eta.sq.part
## sex      0.04292636 0.09409203
## age      0.16266161 0.28242218
## sex:age  0.36537497 0.46923235
```

```
## [1] "No hay efecto medio"
```

```
## [1] "Efecto grande en:age"      "Efecto grande en:sex:age"
```

Analisis Discriminante Lineal (LDA)

Datos

Explorando discriminación por pares de variable

Homogeneidad de la Varianza: Histograma VariablesGrupo

Contraste de Normalidad Univariante Shapiro-Wilk

Contraste de Normalidad MultiVariante

Outliers

Test de Royston

Test de Henze-Zirkler

Contraste de Matriz de Covarianza

Estimación de parámetros de la función de densidad ($\hat{u}(X), E$) y cálculo de la función discriminante según aproximación de Fisher via `lda()`

Evaluación del error: Accuracy Table

Visualización de las clasificaciones

Analisis Discriminante Cuadrático (QDA) > falta de homocedasticidad/outliers LDA

Explorando discriminación por pares de variable

Contraste de Normalidad Univariante Shapiro-Wilk

Contraste de Normalidad MultiVariante

Outliers

Test de Royston

Test de Henze-Zirkler

Contraste de Matriz de Covarianza

Parámetros de la función de densidad función discriminante según aproximación de Fisher via `qda()`

Evaluación del error: Accuracy Table

Visualización de las clasificaciones

Analisis Discriminante Cuadrático Robusto (RQDA) > falta normalidad

Máquinas de Soporte Vectorial

Datos

Grafico datos

Busqueda de mejor hiperparametro C (coste)

Mejor modelo según hiperparametro

3

Predicciones del Modelo

Regresión logística LR